Atomic Radii: Incorporation of Solvation Effects

BRIAN J. SMITH,[1] NATHAN E. HALL[1,2]

[1] *Biomolecular Research Institute, 343 Royal Parade, Parkville, VIC 3052, Australia*

[2] *School of Chemistry, University of Melbourne, Parkville, VIC, Australia*

Received 22 December 1997; accepted 22 April 1998

ABSTRACT: Atomic radii used to define the solute cavity in continuum-based methods are determined by reproducing the solvent-accessible surface defined as the loci of minima in a potential (solvent interaction potential) between the solute and a probe. This potential includes electrostatic interaction (ion−dipole, ion−quadrupole, and ion-induced dipole) terms as well as a Lennard−Jones energy term. The method alleviates the need to distinguish solute atoms in different chemical environments. These radii, when used in the calculation of solvation free energies, are shown to be superior to fixed atom-specific radii or to radii obtained from the electron isodensity surface from quantum-mechanical calculations. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 1482−1493, 1998

Keywords: atomic radii; solute cavity; solvation; electrostatics; solvent interaction potential

Introduction

Continuum (macroscopic) models have become a popular method for describing the effect of solvents on molecular systems, and have been successfully incorporated into quantum-mechanical calculations and simple electrostatic models. A common feature of all these methods is the requirement to determine the solute cavity. The cavity is, however, simply a conceptual simplification; there is no physical boundary between the solute and solvent. Consequently, all sets of radii will depend on the model for which they have been determined, and there cannot be an optimal set of atomic radii for any solute. Nonetheless, these models do provide an efficient and reliable method for determining solvation effects, provided the cavity is of a realistic size and shape.

The Onsager model[1] provides the simplest approach, employing a spherical cavity. This method can be easily incorporated into quantum-mechanical calculations, which makes it particularly attractive.[1−5] Ellipsoidal cavities can provide a more general boundary between the dielectric regions.[6−9] Simple geometrical shapes like the sphere or ellipse are, however, unrealistic for many molecular species. Several methods, such as the polarizable

Correspondence to: B. J. Smith

CCC 0192-8651 / 98 / 131482-12

continuum model (PCM),[10,11] the conductor-like screening model (COSMO),[12–18] and the finite difference Poisson–Boltzmann (FDPB)[19] approach, use atom-centered radii to generate the solute cavity. Many different sources of atomic radii have been considered. Ionic (Bondi)[20] radii determined from the analysis of crystal structures, and radii derived by Pauling,[21] are two such sets that have been used extensively to define the solute boundary. For ionic solutes, it was found that the radii needed to be increased by 0.1 Å for anions and by 0.85 Å for cations to reproduce experimental solvation energies.[22] Subsequently, an empirical factor of 7.0% was chosen for scaling crystallographically derived radii to best fit experimental solvation energies of ionic species.[23]

Standard Pauling radii are often enlarged by 20% in PCM[24,25] calculations. The choice of a scaling factor of 1.2 has been rationalized through molecular dynamics simulations.[26] Implementation of a simple scaling factor of 1.2, however, was found to produce erroneous results for the glycine zwitterion.[27] To overcome this discrepancy, the radii of the NH_3 and O spheres needed to be reduced by 0.4 Å. Further investigations of many systems has revealed that different scaling factors are necessary, depending on the charge of the solute: a scaling factor of 1.10–1.15 was found to be appropriate for ions[28] and a factor of 1.20–1.25 appropriate for neutral species.[29] A further distinction can be applied between polar and nonpolar hydrogens.[29,30]

Gavezzotti has reported an alternative set of atomic radii,[31] based on x-ray crystallographic results, which have been used successfully in FDPB calculations.[32] A set of radii developed by Rashin[33] has also been implemented in a number of solvation studies.[15,34,35] To satisfactorily reproduce experimental solvation energies it was found necessary to distinguish between atoms that are hydrogen bonding and nonhydrogen bonding. Other parameterized sets of atomic radii have been developed specifically for other methods, including the Langevin dipole method of Warshel,[36] and a molecular mechanics method by Bliznyuk and Gready.[37] More complicated and extensive sets of radii, based on the many atom types used in some molecular mechanics force fields, have also been developed, although these radii are largely dependent upon the charge sets employed.[38]

Individual collections of atomic radii, having only a small number of atom types, are generally not sufficiently flexible to account for different atomic environments or charges. With many different atom types, however, difficulties arise when the assignment of atom type becomes ambiguous or inappropriate, particularly in transition states and exotic minima. An approach that eliminates these problems is one in which the atomic radii are determined as a function of the partial atomic charge, although additional factors of basis set dependence and choice of method for determining the charges is introduced where charges are obtained quantum mechanically. Several relationships between partial atomic charge and radii have been determined previously.[39–45] This approach appears general and can be easily applied, although there can still be large errors in the calculation of solvation energies. In the very successful SMx models of Cramer and Truhlar,[46] the atomic radii are a function of the partial atomic charge, which is determined self-consistently.

An alternative to using atomic-centered radii to generate the solute cavity is to use the electron isodensity surface from quantum-mechanical calculations. One of the benefits of such a scheme is that the extensive parameterization that atomic radii methods often require is replaced by a single parameter, the electron density cutoff, typically in the range 0.0004–0.001 a.u. ($e \cdot bohr^{-3}$).[25] Using this approach, however, a new variability is introduced as the isodensity surface is dependent upon the basis set and level of theory used to derive the electron density.

Investigations of the solvation energies of simple neutral, anionic, and cationic species have been performed where the electron density has been used to generate the cavity surface for FDPB calculations.[47] A single cutoff value is found not to be appropriate for all molecular species. Anions, with more diffuse electron distributions, require larger cutoff values than do neutral molecules. Similarly, cations require a different cutoff, again to best reproduce experimental solvation energies. The study of the effect of basis set variation also found that, depending upon the basis set being used, different cutoff values are required. Similar conclusions to these were drawn by Stefanovich[16] who used an isodensity cutoff value of 0.001 a.u. to calculate the solvation energies of neutral and ionic species. The solvation energies of anions were found to be underestimated (resulting from the diffuse electron distribution) and solvation energies of cations overestimated (resulting from the more contracted electron distribution). The assignment of a cutoff value to be used for zwitterionic

systems becomes ambiguous as neither the cutoff applicable for neutral, positive, nor negative species would appear to be appropriate.

It is clear that there is no automatic method for determining the solute cavity that is general enough to deal successfully with all systems (including transition states, neutral, charged, and zwitterionic species) equally well. Problems stem from the need to either define atom types for assignment of parameterized atomic radii or determine appropriate values for the electron density cutoff.

Solvent Interaction Potential

A general method that has been designed to overcome the current deficiencies in determining the solute cavity is described here. The method determines atomic radii (to be applied with FDPB calculations) from a potential established between the solute and a solvent probe, the solvent interaction potential (SIP). The radii, r_{SIP}, are obtained by reproducing the solvent-accessible surface, defined as the loci of minima in the potential.

The solvent interaction potential incorporates both electrostatic and nonelectrostatic interactions between the solute and the probe. The nonelectrostatic (van der Waals) energy is calculated using a Lennard–Jones (LJ) potential, E_{LJ}, while the electrostatic energy is determined from the interaction of the solute partial atomic charges with the dipole, E_{i-d}, quadrupole, E_{i-q}, and ion-induced dipole, E_{i-id}, of the probe. The total solvent interaction potential energy E_{SIP} of the probe with the solute is given by:

$$E_{\mathrm{SIP}} = E_{\mathrm{LJ}} + E_{i-d} + E_{i-q} + E_{i-id} \qquad (1)$$

Here, many factors which also contribute to the total solvation free energy are assumed to cancel.[49] The electrostatic potential of the solute is represented using atomic charges that are determined quantum mechanically, while the probe is assigned the experimental electrostatic properties (dipole, quadrupole, and polarizability) of the solvent. Although the method should be general enough to deal with any solvent, it is demonstrated here using water.

A Lennard–Jones potential, established between the probe and each atom of the solute, is used to represent the van der Waals repulsive and attractive dispersion forces between the solute and the probe. The total LJ energy at any point on the solvent interaction potential is given by the interaction of the probe with each of the solute atoms:

$$E_{\mathrm{LJ}} = \sum_{i}^{N} 4\varepsilon_i \left\{ \left(\frac{\sigma_i}{r} \right)^{12} - \left(\frac{\sigma_i}{r} \right)^{6} \right\} \qquad (2)$$

This LJ potential is defined by two parameters: the well depth of the interaction between the probe and solute atom, ε_i,* and the separation at which the potential is zero, σ_i, (where σ_i is related to the equilibrium bond separation by $2^{-1/6} r_e$). The equilibrium separation, r_e, is equal to the sum of the Lennard–Jones radii, r_{LJ}, and the solvent probe radii, r_{P}. For all interactions between the probe and hydrogen atoms, ε_i is assigned a value of 0.17 kJ mol^{-1}, whereas, for the first-row atoms, C, N, O, and F, a well depth of 0.34 kJ mol^{-1} is used. These values are taken from parameters for He–Ne and Ne–Ne interactions respectively.[49] The r_{LJ}, values need be generated for each individual atom type. The atomic radii determined here depend only on the element type so that no assignment of atoms in different chemical environments needs to be made. This restriction allows for the equal implementation of the solvent interaction potential for all systems over any potential energy surface, including exotic minima and transition states. Consequently, different charged states, including neutral, cationic, anionic and zwitterionic systems and charge transfer processes are all dealt with in an identical manner.

The experimental gas-phase dipole moment, μ, and quadrupole moment tensors, Q_{aa}, of the solvent are scaled by a Langevin function, L, to provide the average effective moments (dipole μ', and quadrupole Q'_{aa}) in the direction of the field generated by the solute partial atomic charges (F) when it is subject to electrical orientation and thermal randomizing forces:

$$\mu' = L\mu \qquad (3)$$

$$Q'_{aa} = LQ_{aa} \qquad (a = x, y, z) \qquad (4)$$

where:

$$L = \frac{e^x + e^{-x}}{e^x - e^{-x}} - \frac{1}{x} \quad \text{and} \quad x = \mu F/\mathrm{kT}$$

Under the influence of low electric fields, the Langevin function tends toward zero, corresponding to totally random orientations of the solvent dipole.

* The well depth, ε_i, should not be confused with the dielectric constant of either the solute or solvent.

The Langevin dipole of the probe is assumed to lie in the opposite direction to the electrostatic potential produced by the solute partial atomic charges, q_i. At any point about the solute, each atom of the solute makes an angle θ_i with the probe dipole, at a distance of r_i. Based on the assumption that the quadrupole is able to freely rotate about its principle axis, the total dipole and quadrupole interaction energies are given by the following expressions[50]:

$$E_{i-d} = \sum_i^N \frac{-\mu' q_i}{\varepsilon r_i^2} \cos\theta_i \tag{5}$$

$$E_{i-q} = \frac{1}{2}\sum_i^N \left[\frac{Q'_{zz} q_i}{\varepsilon r_i^3}(3\cos^2\theta_i - 1)\right] \tag{6}$$

The experimental dipole value of water, 1.86 Debye,[51] and the experimental quadrupole tensor, $Q_{zz} = -0.13$ Buckinghams,[51] have been used here.

A polarizable body situated in the electric field arising from a charge will have induced in it a small dipole. The interaction of the charge with this induced dipole leads to an additional interaction energy. Again, assuming free rotation about the molecular axis (z), the equation for the ion-induced component of the solvent interaction potential is given by:

$$E_{i-id} = \frac{1}{2}\sum_i^N \left[\frac{q_i^2}{r_i^4 \varepsilon^2}\left(\alpha_{zz}\cos^2\theta_i + \frac{(\alpha_{xx} + \alpha_{yy})}{2}\sin^2\theta_i\right)\right] \tag{7}$$

The Langevin function is not applied to the polarizability components α_{xx}, α_{yy}, and α_{zz}, as thermal motions do not effect the ion-induced dipole interaction energy between the solute and the probe. The experimental polarizabilities of water used here are $\alpha_{xx} = 1.528 \times 10^{-24}$, $\alpha_{yy} = 1.415 \times 10^{-24}$, and $\alpha_{zz} = 1.468 \times 10^{-24}$ cm^3.[51]

Solvent-Accessible Surface

The solvent-accessible surface (SAS) is determined from the loci of minima of the solvent interaction potential (SIP), evaluated over a grid surrounding the solute; a regular grid spacing of 0.1 Å has been applied here. Because the center of mass and the probe center do not necessarily coincide, an offset (Γ) is applied in the direction of the electric field. Thermal motions are incorporated using the Langevin of the offset value:

$$\Gamma' = L\Gamma \tag{8}$$

When the solute atomic charges create only a small field, the offset will tend to zero. Conversely, the offset will be greatest under the influence of a large electric field. The effect of the offset under the influence of a positive charge is to expand the SAS, while the SAS is contracted near a negative charge. To determine whether a grid point is part of the SAS, the solvent interaction energy (E_{SIP}) must project a minimum along the direction to the nearest atom. This analysis is performed at a distance from the grid point given by the Langevin offset, but applied in the opposite direction to the electrostatic field (Fig. 1).

Atomic Radii from Solvent-Accessible Surface

Atomic radii, r_{SIP}, are obtained from the solvent-accessible surface generated from the solvent interaction potential (SAS_{SIP}) by determining the radii that best reproduce this surface. Grid points, which lie within the calculated SAS (each representing a cubic volume element), yield the solvent-excluded volume, SEV_{SIP}. Initial estimates of

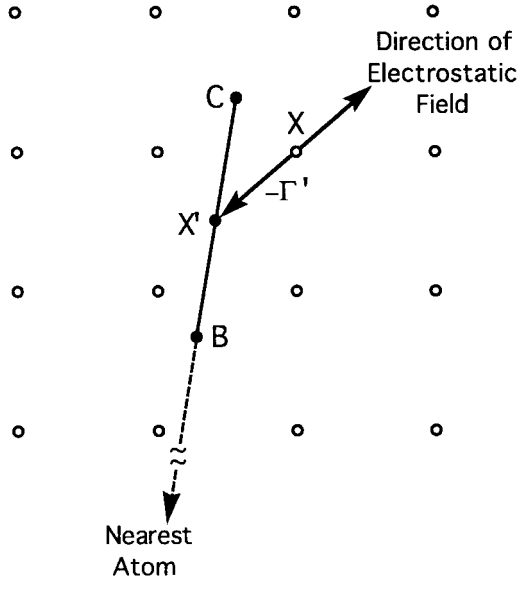


FIGURE 1. The grid point X is part of the SAS if the energy at the point X' projects a minimum along the direction to the nearest atom; the interaction energy at X' must be less than the energy at points B and C. Points B and C are 0.1 Å (single grid spacing) from X'.

the radii are used to generate a representative solvent-accessible surface, SAS_{rep}, and the corresponding solvent-excluded volume, SEV_{rep}. The mismatch in the two volumes is given by those elements that do not belong to both the SEV_{SIP} and SEV_{rep}. Each atomic radius is adjusted in turn until the sum of the mismatch elements is minimized, leading to the optimized r_{SIP} values.[52]

An estimate of the surface area, A_{SIP}, defined by the SAS_{SIP}, is obtained from the sum of the area of the exposed faces of the volume elements. An exposed face is defined as a face shared between two volume elements of which one belongs to the SEV_{SIP} and the other the solvent.

Calculation of Solvation Free Energies

The electrostatic contribution to the solvation free energy is calculated here using the FDPB approach. The radii generated from the SIP, r_{SIP}, are applied in conjunction with the atomic charges used to define the SIP of the solute. Calculations of the nonelectrostatic component of solvation, incorporating the cavitation and dispersion terms, are performed using the relationship based on the solvent-accessible surface area:

$$\Delta G_{\text{non-elec}} = \gamma A_{SIP} + b \tag{8}$$

Values of the constants γ and b, 20.0 J mol^{-1} Å^{-2} and 3.5 kJ mol^{-1}, respectively, are taken from Sitkoff et al.[53] The total solvation free energy is given by the sum of the electrostatic and nonelectrostatic components:

$$\Delta G_{solv} = \Delta G_{elec} + \Delta G_{\text{non-elec}} \tag{9}$$

It is worth noting that, although the electrostatic component of the solvation energy is size consistent, the nonelectrostatic component is not.

Computational Details

Standard molecular orbital calculations[54] have been performed using the Gaussian-94 program.[55] Geometries of the solutes have been optimized at the MP2/6-31G(d) level. CHELPG charges[56] calculated at the HF/6-31+G(d) level are used to define the atomic charges of the solute. CHELPG charges are preferred as they are generated from the electrostatic potential rather than Mulliken charges, which are derived from the arbitrary partitioning of electron density of the electronic wave function.

FDPB calculations are performed using the DelPhi program[57] with an internal solute dielectric of 1 and an external dielectric of 78.54. A grid extent of 25 Å with 0.25 Å resolution is used in these calculations. The CHELPG HF/6-31+G(d) charges used in the FDPB calculations are those used to define the SIP. Calculation of atomic radii from the SIP takes a comparable amount of time as the FDPB calculations.[†]

Parameterization

The Lennard–Jones component of the SIP requires the van der Waals radii of each atom. In addition, an appropriate offset needs to be determined. For the atoms H, C, N, O, and F, r_{LJ} that lead to r_{SIP} are determined so as to best reproduce the solvation free energies of a set of solute molecules comprising the first-row hydrides and methylhydrides, and their anions and cations. These systems are chosen as their experimental solvation free energies are relatively well determined. The solvation energies of CH_3^- and NH_2^- are unreliable, and therefore not included in this set. The aim of this procedure is to calculate solvation free energies within an accuracy of 10 kJ mol^{-1} of experimental solvation free energies.

The difficulties in determining accurate solvation energies of ions are well recognized, with uncertainties in pK_a values and gas-phase acidities and basicities being the main sources of error. Solvation energies for ions have been reevaluated here using recent experimental gas-phase acidities and basicities.[58] These are presented in Table I. The solvation energies obtained are very similar to those recently obtained by Florián and Warshel.[63] Experimental ΔG_{solv} values of the neutral species are based on values from Hine et al.[59] and Wagman et al.[60] with standard state corrections.[61]

The optimized Lennard–Jones radii (r_{LJ}) and offset (Γ) for the parameterization set of molecules are as follows (in Å); 0.17 for the probe offset and r_{LJ} of C = 1.90, N = 2.10, O = 1.80, F = 1.72, and H = 0.80. The average error in the solvation free energies calculated for the parameterization set using the SIP method is −0.3 kJ mol^{-1}, with an average unsigned error of 4.1 kJ mol^{-1}. These results are presented in Table II. All calculated

[†] These calculations require only a few minutes on a SGI R10000 for the systems presented here.

TABLE I.
Evaluation of Experimental Free Energies of Solvation of Ions.[a]

	$\Delta G_{acid/base}$[b]	$\Delta G^0_{solv}(HX)$[c]	pK_a[d]	ΔG^0_{solv}
OH^-	1607.0	−26.4	15.74	−458.0
CH_3O^-	1567.0	−21.3	15.5	−414.3
F^-	1530.5	−31.4	3.18	−458.0
NH_4^+	−817.1	−18.0	9.27	−339.4
$CH_3NH_3^+$	−867.1	−19.2	10.62	−298.3
H_3O^+	−660.9	−26.4	−1.75	−441.3
$CH_3OH_2^+$	−729.6	−21.3	−3.0	−360.4

[a] Experimental free energies of solvation of ions are determined using the following expressions:

$$\Delta G^0_{solv}(X^-) = -\Delta G_{acid}(HX) - \Delta G^0_{solv}(H^+) + \Delta G^0_{solv}(HX) + 5.69\ pK_a \quad \text{(anions, } X^-\text{)}$$

$$\Delta G^0_{solv}(BH^+) = -\Delta G_{base}(BH^+) + \Delta G^0_{solv}(H^+) + \Delta G^0_{solv}(B) - 5.69\ pK_a \quad \text{(cations, } BH^+\text{)}$$

where ΔG_{acid} is the gas-phase acidity of HX and ΔG_{base} the gas-phase basicity of B, and ΔG^0_{solv} the solvation free energies of the proton or parent base, HX or B. The solvation energy of the proton is taken as -1085.8 kJ mol^{-1} (units are in kJ mol^{-1}).
[b] From ref. 58.
[c] From refs. 59 and 60, with standard state corrections.[61]
[d] From ref. 62.

solvation energies lie within 10 kJ mol^{-1} of the experimental values, including the methoxide anion, which has generally proven difficult to reproduce via computational methods.[32, 64, 65] Because there is no accurate experimental solvation energies for nitrogen anions (NH_2^- or CH_3NH^-) parameterization of the nitrogen atom has been applied to neutral and cationic species only. Thus, while the solvation energies for the nitrogen species are well reproduced, the parameterization may be

TABLE II.
Comparison of Calculated Solvation Free Energies (kJ mol^{-1}).[a]

		FDPB[b]					
	ΔG_{expt}	r_{SIP}	r_A	r_ρ	AM1-SM2[c]	ILD[d]	PCM[e]
CH_4	9.2	−2.8	−2.9	−1.7	−5.9	−1.7	−9.2
CH_3CH_3	7.5	0.6	0.6	2.0	2.5	0.9	−7.5
NH_3	−18.0	5.5	5.6	0.8	0.0	1.3	−6.0
CH_3NH_2	−19.2	8.4	10.4	7.2	−6.5	7.1	−3.5
H_2O	−26.4	−9.2	−13.3	−5.6	0.0	−10.4	4.2
CH_3OH	−21.3	−4.4	−5.3	1.8	−3.0	−3.8	−6.2
HF	−31.4	2.0	−0.8	−0.6	31.9	10.9	11.1
CH_3F	−0.8	−6.8	−9.6	−6.1	2.5	−10.9	−18.0
OH^-	−458.0	−3.9	2.4	85.5	5.7	−23.2	−15.6
CH_3O^-	−416.7	8.8	43.4	94.4	67.3	10.9	−13.4
F^-	−458.0	1.1	1.1	62.4	10.3	22.9	−12.6
NH_4^+	−337.1	−4.2	−0.5	−48.6	6.7	−1.8	−39.7
$CH_3NH_3^+$	−300.9	−1.5	−2.9	−24.0	0.1	−8.7	−34.7
H_3O^+	−440.9	1.7	1.2	−15.6	6.6	18.3	30.8
$CH_3OH_2^+$	−365.5	5.8	6.8	3.7	13.2	18.2	17.1

[a] Results for each of the computational methods are presented as differences (theory minus experiment).
[b] Finite difference poisson–Boltzmann method using SIP derived radii (r_{SIP}), fixed atom-specific radii (r_A), and radii from the isodensity surface (r_ρ).
[c] From ref. 64.
[d] From ref. 63.
[e] Calculated at the HF/6-31+G(d) level upon geometries optimized at the MP2/6-31G(d) level. van der Waals radii (H:1.2, C:1.5, N:1.5, O:1.4, F:1.35) were unscaled.

slightly biased toward the positive species. This may result in nitrogen having an artificially large radii, but the lack of reliable experimental data leaves no avenue for testing this. Similarly, the carbon radius is parameterized only in a methyl (CH_3) environment, which may have biased the determination of the carbon r_{LJ} value. The parameterization is also generally biased toward the ionic systems, because the solvation energy of ions is more sensitive to changes in the radii than neutral systems.

The largest components of the solvent interaction potential are the dipole interaction and the Lennard–Jones energies, with the magnitudes of the quadrupole and ion-induced dipole energies much smaller by comparison. Presented in Table III are the average energies for each of the individual components of the SIP, which define the SAS (minima in the potential) for H_2O, OH^-, and H_3O^+. Omitting the quadrupole and ion-induced dipole terms in the calculation of the solvent interaction potential gives rise to almost identical r_{SIP} values. The largest variation in radii is 0.02 Å, with most radii remaining unchanged.

Radii determined from the SIP using MP2/6-311 + G(d, p) CHELPG charges are almost identical to those obtained using the HF/6-31 + G(d) CHELPG charges, despite quite large differences (in some cases 0.1 *e*). There appears, therefore, no benefit associated with using more elaborate methods to calculate charges than HF/6-31 + G(d) for these calculations. Only minor changes were observed when MP2/6-31 + G(d)-optimized geometries were used in place of MP2/6-31G(d) geometries, showing no requirement for the computationally more expensive geometry optimizations. In general, the solvent-excluded volume generated by the atom-centered spheres reproduces the volume generated by the SIP to within 1%. Of the systems studied here, the largest deviation in volumes ($SEV_{SIP} - SEV_{rep}$) is 2.2 Å^3 (of a total solvent-excluded volume of 130 Å^3) for the water molecule, a difference of 1.7%. The atomic radii, however, are derived by minimizing the total difference between the SEV_{SIP} and the SEV_{rep}. Thus, some regions of the SEV_{rep} extend beyond the SEV_{SIP}, and some regions of the SEV_{SIP} are not covered by the SEV_{rep}. The total mismatch in volume for water, for example, is 6.2 Å^3; 2.0 Å^3 exists in the SEV_{SIP}, which the SEV_{rep} does not cover, and the SEV_{rep} extends beyond the SEV_{SIP} by 4.2 Å^3. The total mismatch in volumes does not exceed 5% of the SEV_{SIP} for any of the systems here.

TABLE III.
Components of Total Solvent Interaction Potential (kJ mol^{-1}).

	E_{LJ}	E_{i-d}	E_{i-q}	E_{i-id}
H_2O	−0.410	−0.138	0.000	−0.002
OH^-	−0.181	−0.785	0.018	−0.004
H_3O^+	−0.480	−0.646	−0.014	−0.002

To assess the effect of the SIP method a set of fixed atom-specific radii, r_A, which best reproduced the solvation free energies of the parameterization set of solutes, were determined. The same solute geometries, CHELPG charges, and FDPB conditions were used for these calculations to enable a direct comparison of solvation energies. The optimized atomic radii (r_A) were found to be (in Å) H = 1.04, C = 1.69, N = 2.12, O = 1.57, and F = 1.48. The errors in the calculated solvation free energies using these radii are presented in Table II. These radii are generally quite similar to those of Bondi,[20] with the exception of hydrogen and nitrogen atoms, which are considerably smaller (0.16 Å) and larger (0.57 Å), respectively. For neutral species, the solvation free energies using r_A values compare quite favorably with the energies obtained through r_{SIP}. Errors in solvation free energies are found to be only slightly larger using r_A radii, although the largest value for water, 13 kJ mol^{-1}, is larger than the target value of 10 kJ mol^{-1}. For the anions, the calculated solvation free energies of OH^- and F^- are in good agreement with experiment. The error for methoxide, however, is almost 45 kJ mol^{-1}. The cationic species all perform well, with no errors in excess of 10 kJ mol^{-1}. The average and unsigned-average errors using these r_A radii are 2.4 and 7.1 kJ mol^{-1}, respectively, significantly larger than those from the SIP method.

Atomic radii can also be obtained from the electron isodensity surface of quantum-mechanical calculations, r_ρ. In this case the only variable is the cutoff. A cutoff of 0.0017 a.u. was found to minimize the total (unsigned) error in the calculated solvation free energies of the set of molecules in the parameterization set. Electron densities were calculated over a regular (0.1 Å) cubic grid surrounding the solute molecule at the HF/6-31 + G(d) level. The molecular volume includes all grid points having an electron density that is greater than the cutoff. The solvent-excluded volume includes the molecular volume plus all other grid points within the probe radius of the molecular

volume. Radii were obtained from the solvent-excluded volume by the same method used to generate radii from the SIP (above). Solvation energies were calculated (FDPB) using these radii and HF/6-31 + G(d) CHELPG charges. The errors in the calculated solvation free energies using these radii are presented in Table II.

Energies calculated for the neutral species using the r_ρ radii all have errors of less than 10 kJ mol^{-1}. The anions and cations, however, are all extremely poorly represented. Apart from $CH_3OH_2^+$, the errors for all the charged species are greater than 10 kJ mol^{-1}, with errors as large as 86 and 94 kJ mol^{-1} for hydroxide and methoxide, respectively. Clearly, a uniform value for the electron density cutoff is not appropriate to generate atomic radii for species of different charge. The unsigned average error using radii from the electron density is 24.0 kJ mol^{-1}, making this a particularly unsatisfactory method for determining radii.

Also presented in Table II are the errors in the calculated solvation free energies from three alternative methods, including AM1-SM2,[64] ILD,[63] and PCM.[24,25] AM1-SM2 does quite well with the exception of HF and CH_3O^-, although it has been parameterized using solvation energies of ions that differ significantly from those used here. ILD does reasonably well for CH_3O^-, but rather poorly for the other anions. PCM performs least well of the alternative methods, although it does much better with anions than FDPB with r_ρ radii. The discrepancy can, in part, be attributed to the outlying charge (portion of the electron density that lies beyond the cavity boundary) that is present in all quantum-mechanical calculations.[13] For this limited set of solutes the SIP approach might be expected to perform better than the other methods since it has been parameterized using these solutes, although AM1-SM2 has been parameterized with CH_3O^- and yet it still performs rather poorly for this system.

SIP Atomic Radii

Radii and CHELPG charges of the set of molecules used in the parameterization of the SIP Lennard–Jones radii are presented in Figure 2. It is apparent from this figure that there is a wide variety of radii predicted for atoms in different environments. Thus, for example, the radius of carbon ranges from 1.62 Å on methoxide to 1.90 Å on methylfluoride and, for hydrogen, values range between 0.63 Å on methoxide and 1.16 Å on ammonia.

Alignment of the probe dipole with the electrostatic field ensures a favorable interaction with the solute charges that will cause a reduction in the radii, especially of charged species (both negative and positive), because the Langevin function tends to reduce severely the magnitude of the effective probe dipole surrounding neutral systems. The effect of the offset, however, depends on the local electrostatic field. Generally, the electrostatic field increases in the direction toward a negative charge, and decreases in the direction toward a positive charge. In regions where the electrostatic potential increases in the direction toward the solute, the SAS is moved inward, reducing the radii of nearby atoms. Conversely, the SAS is moved outward, increasing the radii of nearby atoms, where the electrostatic potential decreases in the direction toward the solute. The Lennard–Jones interaction energy between the probe and hydrogen atoms is much weaker than the interaction with oxygen. As a result, the radii of hydrogen atoms are largely dependent on the atoms to which they are bonded. This produces the large radii for hydrogen atoms bonded to nitrogen.

The effects of the offset and dipole interaction is best illustrated through inspection of the trend in radii of the series H_3O^+, H_2O, and OH^-. Without an offset being applied, the oxygen atom radius in H_2O is found to be 1.76 Å, just 0.03 Å shorter than the Lennard–Jones radius for the oxygen atom. The oxygen atom radii in H_3O^+ and OH^- are substantially smaller, 1.67 and 1.71 Å, respectively, as a result of stronger ion–dipole interactions. The offset decreases the oxygen atom radius in H_2O by 0.07 Å and increases the hydrogen atom radius by 0.20 Å. The oxygen atom radius in H_3O^+ actually increases by 0.16 Å despite the oxygen atom carrying a negative charge; the electrostatic potential decreases in the direction of the solute at all points surrounding the solute. Thus, the direction in which the offset is applied does not necessarily correspond to the charge of the nearest atom. In OH^- the oxygen atom radius is quite small, resulting from the cooperative effects of both favorable ion–dipole interactions and the probe offset.

Radii from Electron Density

The hydrogen atom radii obtained from the electron isodensity surface are generally larger than the radii from the SIP method. They range from

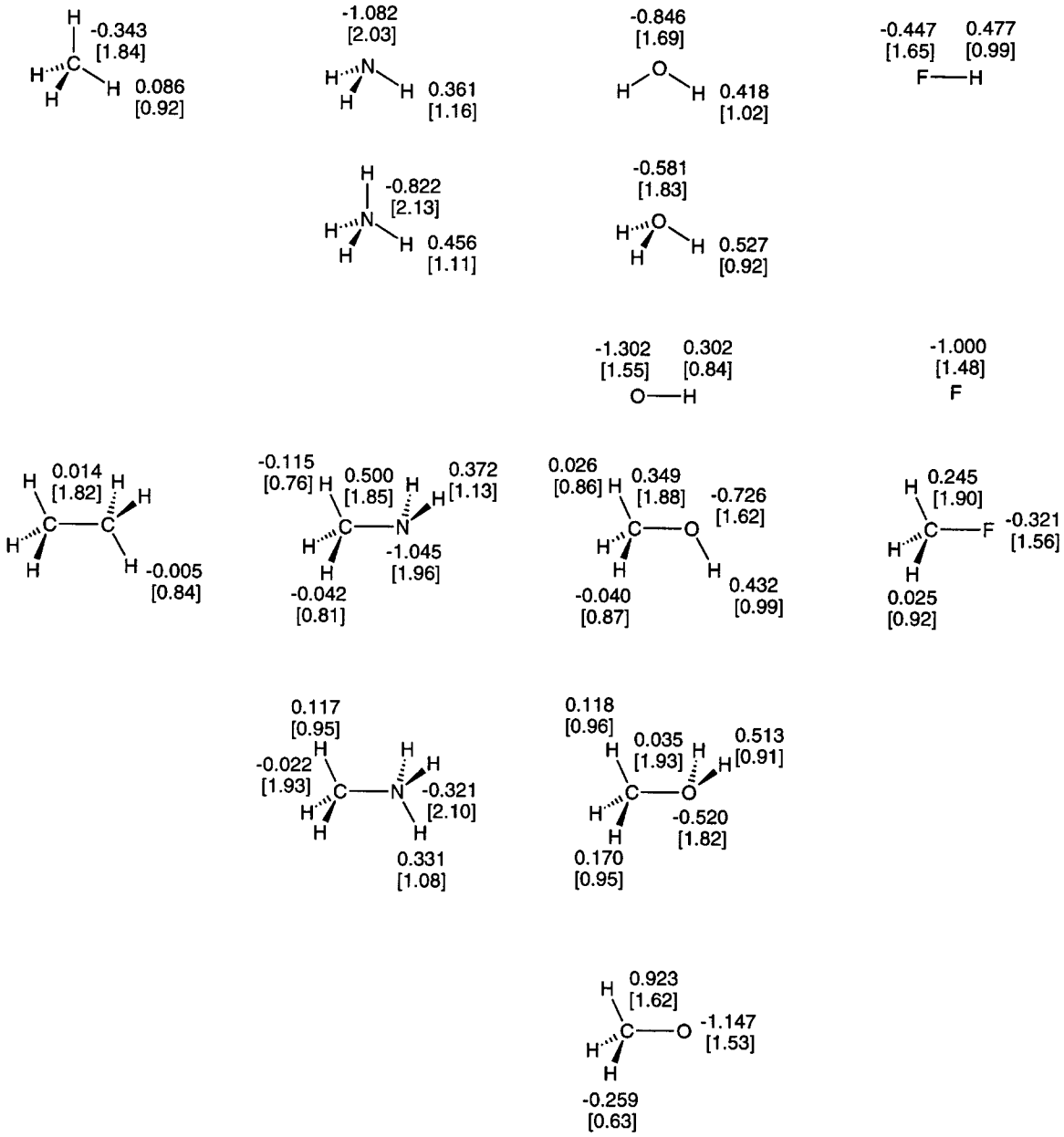


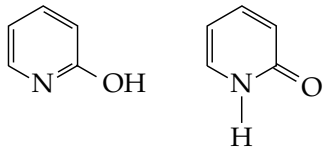
FIGURE 2. Calculated HF / 6-31+G(d) CHELPG charges and SIP atomic radii (in square brackets).

0.96 Å in the hydronium ion to 1.39 Å in methoxide. This is in accordance with radii determined previously from electron densities[66] and the expectation that anions possess a more diffuse electron distribution than neutral and cationic species. Atomic radii decrease along the CH_4, NH_3, H_2O, HF series for the non-hydrogen atom. The carbon atom radius in methane is 1.91 Å, nitrogen in ammonia is 1.85 Å, oxygen in water is 1.72 Å, and fluorine in HF is 1.57 Å. The hydrogen atom radii also decrease along this series.

Applications of Radii from SIP

The differential free energy of solvation of the tautomer equilibrium between 2-hydroxypyridine (**A**) and 2-pyridone (**B**) has been measured experimentally in water.[67] The observed difference, 18.0 kJ mol^{-1}, is well reproduced by the semiempirical methods, AM1-SM1 and PM3-SM3 (18 kJ mol^{-1}), but underestimated by AM1-SM2 (11 kJ mol^{-1}).[46] Calculations using radii obtained from the SIP

yield a differential solvation energy of 22.6 kJ mol^{-1}, also in good agreement with experiment. The nitrogen radius in **A** (1.82 Å) is significantly smaller than in **B** (1.99 Å). The oxygen radius in **A** is identical to the radius found in methanol, whereas the radius in **B** is 0.06 Å smaller. The N-bonded hydrogen in **B** has a much smaller radius (1.01 Å) than in methanamine (1.13 Å):



A **B**

The experimental solvation energies of the *E* and *Z* isomers of methylacetamide are roughly equal, 41.8 kJ mol^{-1}, very similar to that found for acetamide, 40.6 mol^{-1}.[68] FDPB calculations using charges and radii from the OPLS force field[69] underestimate the solvation energy of the *E* isomer (25.9 kJ mol^{-1}), and overestimate the solvation energy of the unsubstituted amide (54.4 kJ mol^{-1}), but reproduce quite well the solvation energy of the *Z* isomer (41.0 kJ mol^{-1}).[70] Using radii and charges from the SIP, the solvation energies of the *E* and *Z* isomers are predicted to be 42.8 and 39.0 kJ mol^{-1}, respectively, in quite good agreement with experiment. For acetamide, however, the calculated solvation energy is 47.4 kJ mol^{-1}, still somewhat larger than experiment (yet within the 10 kJ mol^{-1} target). The solvation energies obtained using the radii obtained from the SIP method are very similar to the results from AM1-SM2.[46]

The solvation free energies for the methyl amines and their cations are generally rather poorly reproduced using continuum-based methods.[63,71] Solvation energies obtained using radii and charges from the SIP method are no exception to this. Ammonia, methanamine, and their protonated cations were used in determining the Lennard–Jones radii of the SIP, and good agreement with experiment is guaranteed. The solvation free energies for $(CH_3)_2NH$ and $(CH_3)_3N$, however, are calculated to be too small (by 13.4 and 12.1 kJ mol^{-1}, respectively), whereas for $(CH_3)_2NH_2^+$ and $(CH_3)_3NH^+$ the calculated solvation energies, −283.3 and −264.4 kJ mol^{-1}, are considerably greater than experimental estimates, −268.6 and −236.9 kJ mol^{-1}, respectively.[71] Increasing the radii of the cations to improve the agreement with experiment can only make the comparison for the neutral species worse.

Conclusions

A new method, based on a solvation interaction potential, has been successfully developed to determine atomic radii for use in FDPB calculations of solvation free energies. The method is designed to be general and applicable at any point on a potential energy surface, including exotic minima and transition states, irrespective of the total charge. The method has been shown to yield radii that produce solvation free energies of a parameterization set of molecules more accurately than radii obtained from more conventional approaches, such as fixed atom-specific radii or from the electron isodensity surface.

References

1. M. W. Wong, M. J. Frisch, and K. B. Wiberg, J. Am. Chem. Soc., **113**, 4776 (1991).
2. M. W. Wong, K. B. Wiberg, and M. J. Frisch, J. Am. Chem. Soc., **114**, 523 (1992).
3. M. W. Wong, K. B. Wiberg, and M. J. Frisch, J. Am. Chem. Soc., **114**, 1645 (1992).
4. M. W. Wong, K. B. Wiberg, and M. J. Frisch, J. Chem. Phys., **95**, 8991 (1991).
5. K. B. Wiberg and M. W. Wong, J. Am. Chem. Soc., **115**, 1078 (1993).
6. D. Rinaldi, M. F. Ruitz-Lopez, and J.-L. Rivail, J. Chem. Phys., **78**, 834 (1983).
7. D. Rinaldi, J.-L. Rivail, and N. Rguini, J. Comput. Chem., **13**, 675 (1992).
8. C. Chipot, D. Rinaldi, and J.-L. Rivail, Chem. Phys. Lett., **191**, 287 (1992).
9. G. P. Ford and B. Wang, J. Comput. Chem., **13**, 229 (1992).
10. J. Tomasi and M. Perisco, Chem. Rev., **94**, 2027 (1994).
11. S. Miertus, E. Scrocco, and J. Tomasi, Chem. Phys., **55**, 117 (1981).
12. A. Klamt and G. Schüürmann, J. Chem. Soc. Perkin Trans. II, 799 (1993).
13. A. Klamt and V. Jonas, J. Chem. Phys., **105**, 9972 (1996).
14. J. Andzelm, C. Kölmel, and A. Klamt, J. Chem. Phys., **103**, 9312 (1995).
15. T. N. Truong and E. V. Stefanovich, Chem. Phys. Lett., **240**, 253 (1995).

16. E. V. Stefanovich and T. N. Truong, Chem. Phys. Lett., **244**, 65 (1995).
17. E. V. Stefanovich and T. N. Truong, J. Chem. Phys., **105**, 2961 (1996).
18. K. Baldridge and A. Klamt, J. Chem. Phys., **106**, 6622 (1997).
19. See A. R. Leach, Molecular Modelling. Principles and Applications, Longman, New York, 1996.
20. A Bondi, J. Phys. Chem., **68**, 441 (1964). Atomic radii are H = 1.2, C = 1.70, N = 1.55, O = 1.52, and F = 1.47 Å.
21. L. Pauling, The Nature of the Chemical Bond, Cornell University Press, Ithaca, NY, 1960.
22. W. M. Latimer, K. S. Pitzer, and C. M. Slansky, J. Chem. Phys., **7**, 108 (1939).
23. A. A. Rashin and B. J. Honig, J. Phys. Chem., **89**, 5588 (1985).
24. R. Cammi, M. Cossi, and J. Tomasi, J. Chem. Phys., **104**, 4611 (1996).
25. J. B. Foresman, T. A. Keith, K. B. Wiberg, J. Snoonian, and M. J. Frisch, J. Phys. Chem., **100**, 16098 (1996).
26. F. J. Luque, M. J. Negre, and M. Orozco, J. Phys. Chem., **97**, 4386 (1993).
27. R. Bonaccorsi, P. Palla, and J. Tomasi, J. Am. Chem. Soc., **106**, 1945 (1984).
28. M. Orozco and F. J. Luque, Chem. Phys., **182**, 237 (1994).
29. M. Bachs, F. J. Luque, and M. Orozco, J. Comput. Chem., **15**, 446 (1994).
30. M. Cossi, V. Barone, R. Cammi, and J. Tomasi, Chem. Phys. Lett., **225**, 327 (1996).
31. A. Gavezzotti, J. Am. Chem. Soc., **105**, 5220 (1983).
32. I. Alkorta, H. O. Villar, and J. Perez, J. Comput. Chem., **14**, 620 (1993).
33. A. A. Rashin and K. Namboodiri, J. Phys. Chem., **91**, 6003 (1987).
34. A. A. Rashin, M. A. Bukatin, J. Andzelm, and A. T. Hagler, Biophys. Chem., **51**, 375 (1994).
35. T. Furuki, A. Umeda, M. Sakurai, Y. Inoque, and R. Chûjô, J. Comput. Chem., **15**, 90 (1994).
36. V. Luzhkov and A. Warshel, J. Comput. Chem., **13**, 199 (1991).
37. A. A. Bliznyuk and J. E. Gready, J. Phys. Chem., **99**, 14506 (1995).
38. D. Horvath, D. van Belle, G. Lippens, and S. J. Wodak, J. Chem. Phys., **104**, 6679 (1996).
39. M. A. Aguilar and F. J. Olivares del Valle, Chem. Phys., **129**, 439 (1989).
40. F. J. Olivares del Valle and M. A. Anguilar, J. Comput. Chem., **13**, 115 (1992).
41. F. J. Olivares del Valle, Chem. Phys., **170**, 161 (1993).
42. B. Roux, H.-A. Yu, and M. Karplus, J. Phys. Chem., **94**, 4683 (1990).
43. S. W. Rick and B. J. Berne, J. Am. Chem. Soc., **116**, 3949 (1994).
44. S. Miertus, J. Bratos, and M. Trebatická, J. Mol. Liquids, **33**, 139 (1987).
45. O. Takahashi, H. Sawahata, Y. Ogawa, and O. Kikuchi, J. Mol. Struct. (Theochem), **393**, 141 (1997).
46. C. J. Cramer and D. G. Truhlar, In Reviews in Computational Chemistry, *Vol. VI*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1995.
47. C. Lim, S. L. Chan, and P. Tole, In Structure and Reactivity in Aqueous Solution, C. J. Cramer and D. G. Truhlar, Eds., ACS, Washington, DC, 1994.
48. R. Bonnaccorsi, C. Ghio, and J. Tomasi, Studies Phys. Theor. Chem., **21**, 407 (1982).
49. G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham, Intermolecular Forces, Clarendon Press, Oxford, 1981.
50. A. D. Buckingham, Disc. Faraday Soc., **24**, 151 (1957). See also J. O'M. Bockris and A. K. M. Reddy, Modern Electrochemistry 1, Plenum, New York, 1970.
51. C. G. Gray and K. E. Gubbins, Theory of Molecular Fluids, Clarendon Press, Oxford, 1984.
52. In the process of minimization, the radii of each atom is incremented in turn until the total mismatch increases from the previous step. Following this, the radii are decremented until the mismatch again increases. This process is repeated with decreasing increments until the desired accuracy is obtained.
53. D. Sitkoff, K. A. Sharp, and B. Honig, J. Phys. Chem., **98**, 1978 (1994).
54. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, Ab Initio Molecular Orbital Theory, John Wiley & Sons, 1986.
55. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, J. D. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, *Gaussian-94, Revision C.3*, Gaussian, Inc., Pittsburgh, PA, 1995.
56. C. M. Breneman and K. B. Wiberg, J. Comput. Chem., **11**, 361 (1990).
57. *DelPhi*, *Version 2.50*, Biosym Technologies, San Diego, CA, 1993.
58. (a) J. E. Bartmess, In NIST Standard Reference Database Number 69, W. G. Mallard and P. J. Linstrom, Eds., National Institute of Standards and Technology, Gaithersburg, MD, 20899 (August 1997). (b) E. P. Hunter and S. G. Lias, In NIST Standard Reference Database Number 69, W. G. Mallard and P. Linstrom, Eds., National Institute of Standards and Technology, Gaithersburg, MD, 20899 (August 1997) (http://webbook.nist.gov).
59. J. Hine and P. K. Mookerjee, J. Org. Chem., **40**, 292 (1975).
60. D. D. Wagman, V. B. P. Parker, R. H. Schumm, I. Halow, S. M. Bailey, K. L. Churney, and R. L. Nuttall, J. Phys. Chem. Ref. Data, **11**(Suppl. 2) (1982).
61. C. Lim, D. Bashford, and M. Karplus, J. Phys. Chem., **95**, 5610 (1991).
62. A. Albert and E. P. Sergeant, Ionization Constants of Acids and Bases, Metheun, London, 1962.
63. J. Florián and A. Warshel, J. Phys. Chem. B, **101**, 5583 (1997).
64. C. J. Cramer and D. G. Truhlar, J. Comput.-Aid Mol. Des., **6**, 629 (1992).

65. J. Chen, L. Noodleman, D. A. Case, and D. Bashford, J. Phys. Chem., **98**, 11059 (1994).

66. R. F. Hout Jr. and W. J. Hehre, J. Am. Chem. Soc., **105**, 3628 (1983).

67. P. Beck, Acc. Chem. Res., **10**, 186 (1977).

68. R. Wolfenden, Biochemistry, **17**, 210 (1978).

69. W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc., **110**, 167 (1988).

70. A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T. F. Hendrickson, and W. C. Still, J. Am. Chem. Soc., **113**, 1454 (1991).

71. I. Tuñón, E. Silla, and J. Tomasi, J. Chem. Phys., **96**, 9043 (1992).

72. Experimental solvation energies for $(CH_3)_2NH_2^+$ and $(CH_3)_3NH^+$ were determined from the gas-phase basicities taken from ref. 59 (896.5 and 918.1 kJ mol^{-1}), neutral solvation free energies from refs. 60–62 (18.0 and 13.4 kJ mol^{-1}), and pK_as from ref. 63 (10.77 and 9.8, respectively).